

Supplementary Figure 1. Meta-analysis of overall survival of breast cancer patients carrying the ES signature. We employed the random effects procedure of DerSimonian and Laird (Ref. 49) using inverse variance weighting of hazard ratios estimated from study-specific Cox proportional hazards models. This analysis yielded a log hazard value of 0.71 (hazard ratio 2.03) (grey diamond), indicating that the mortality rate of patients carrying ES-signature tumors is twice as high as that of the remainder of the patients ($P<0.0001$). Black boxes indicate log hazard ratios of individual studies; box sizes indicate relative weight in the meta-analysis; lines indicate 95% confidence intervals. We repeated this analysis excluding patients with tumors classified as basal-like. This yielded a log hazard ratio of 0.87 at $P=0.002$. Additionally, using Cox inverse variance methods we found that combined ES-positive and grade 3 status is more predictive than grade 3 status alone ($P=0.03$).

a

No Cycling Genes

<i>Gene Sets</i>	Grade			ER						
	1	2	3	+	-	Lum A	Normal	Lum B	Her2	Basal
PRC2 targets	7.6	3.5	19	8.1	15	11	9.5	6.4	5.9	
Suz12 targets	9.1	3.6	20	5.3	9.7	11	12	3.5	5.2	5.2
Eed targets	9.1		14	4.9	6.6	9.2	9.9	2.6	6.8	4.2
H3K27 bound	8.7	3.8	20	8.2	13	12	13	9.4	5.3	
Myc targets1					4.2					
Myc targets2						2.6				
Nanog targets	2.6						2.8			
Oct4 targets	2.2						5			
Sox2 targets					2.2		2.5		2.5	
NOS targets				3	2	5.2	9.5		3.4	
NOS TFs			4.2	4.4	9.2				7.2	
ES exp2			12	5.2	11	8.2	2	2	14	
ES exp1	3.5	2	32	8.5	42	6.5			50	

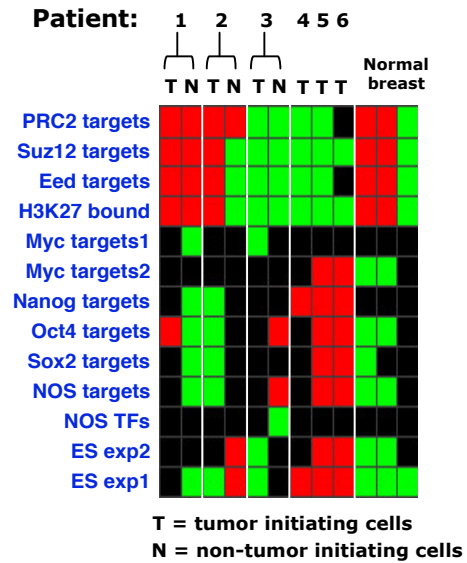
b

No Proliferation Cluster

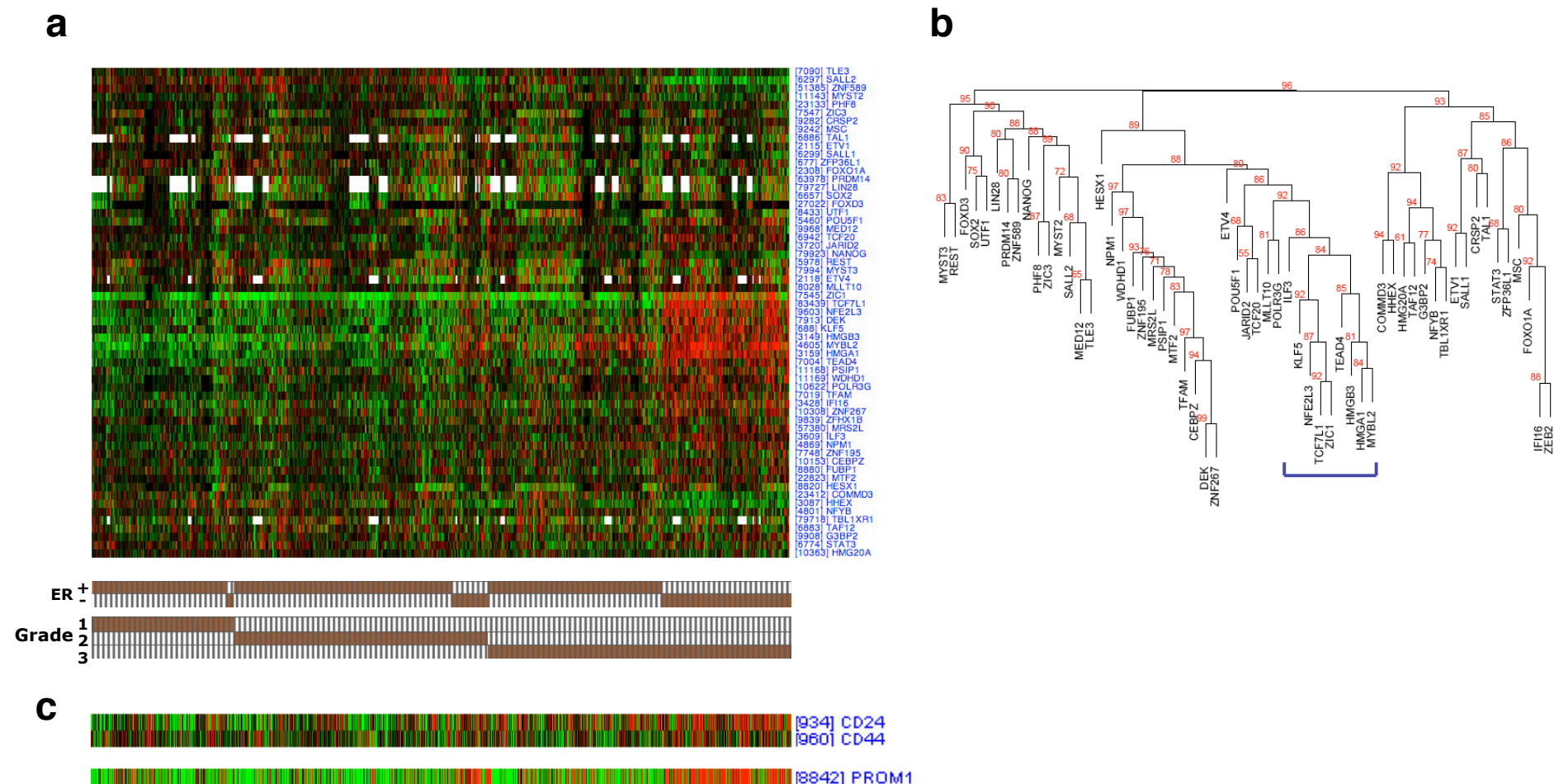
<i>Gene Sets</i>	Grade			ER						
	1	2	3	+	-	Lum A	Normal	Lum B	Her2	Basal
PRC2 targets	6.8	4.1	20	8.1	15	13	9.8	7.5	8.2	
Suz12 targets	9.5	3	17	5.1	9.7	11	12	3.6	5.3	5.1
Eed targets	9.1		14	4.9	6.6	9.6	10	2.4	6.4	4.9
H3K27 bound	8.7	4.7	22	8.7	14	12	12	8.3	6.6	
Myc targets1	2.9		4.9	2.1	6.9	2.6			3.3	
Myc targets2						3.2				
Nanog targets							2.7			
Oct4 targets							4			
Sox2 targets		2.4	2.5						2.3	
NOS targets				2.7	3.2		4.2	6.4	5.9	
NOS TFs				5.5	4.9	2.1			3.7	
ES exp2		2.8	19	7.4	21	14	2.2	2.1	23	
ES exp1	10	3.7	42	15	43	16	3.2		48	

Supplementary Figure 2. Gene-set enrichments in breast compendium tumors following subtraction of proliferation gene sets and of ER status classifiers. (a) Gene sets subtracted for Cycling Genes. (b) Gene sets subtracted for Proliferation Cluster genes.

Breast tumor-initiating cells:



Supplementary Figure 3. Gene-set enrichments in breast cancer tumor-initiating and non-tumor-initiating cell fractions. Gene set enrichments in the tumor-initiating (CD44^{high}/CD24^{low}) fraction (T) and the non-tumor initiating fraction (N) isolated from 3 individual breast tumors (1,2,3) (Ref. 43). Also included are the tumorigenic fraction from 3 additional tumors (4,5,6) and three normal breast samples profiled in this study.



Supplementary Figure 4. ES-associated transcription regulators in breast cancers. (a) Gene expression of 59 ES-associated transcription regulators in breast compendium tumors as in Fig. 6a, with gene symbols added. (b) Hierarchical clustering of the 59 genes using the pvclust R tool (Ref. 50), with 10,000 bootstrap iterations. Scores in red represent statistical significance of each dendrogram node on a 1 to 100 scale. Blue frame denotes cluster of 9 genes highly expressed in grade 3/ER-negative tumors, and defined as the Core 9 set. (c) Expression of cancer stem-cell markers: CD24, CD44, CD133 (PROM1) in the same samples.